# REPORT DOCUMENTATION PAGE

Form Approved
OMB NO. 0704-0188



1. AGENCY USE ONLY ( Leave Blank)

2. REPORT DATE 03/15/02

3. REPORT TYPE AND DATES COVERED
Final, ~~03/01/1998-12/31/2001~~
20 Mar 98 - 19 Dec 01

4. TITLE AND SUBTITLE
Adaptive Control of Nonlinearly Parameterized Systems

5. FUNDING NUMBERS
DAAG55-98-1-0235

6. AUTHOR(S)
Anuradha M. Annaswamy

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Massachusetts Institute of Technology
Department of Mechanical Engineering
77, Mass Ave., Cambridge, MA 02139

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U. S. Army Research Office
P.O. Box 12211
Research Triangle Park, NC 27709-2211

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

36438.5-CI

11. SUPPLEMENTARY NOTES
The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision, unless so designated by other documentation.

12 a. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for public release; distribution unlimited.

12 b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

This project pertains to the development of a theory for adaptive control of nonlinear dynamic systems that are nonlinearly parameterized (NLP). Developments in NLP systems that have been carried out as a part of this project relax the ubiquitous assumption made in the context of adaptive control which is that the unknown parameters occur linearly.
During the past year, we have derived several new results related to NLP systems, and can be grouped under two categories:
(i) Control of nonlinear systems with a triangular structure,
(ii) Parameter convergence in NLP systems.
The class of systems considered in (i) includes high-dimensional nonlinear systems connected in chain and triangular forms, examples of which include Hammerstein-Uryson models and recurrent neural networks. Global stabilization and tracking can be guaranteed for such systems in the presence of unknown parameters that occur nonlinearly.
The results related to (ii) pertain to conditions of *persistent excitation* (PE) under which parameter convergence occurs in a class of discrete and continuous NLP systems. It is shown that for different classes of transcendental functions, distinct PE conditions can be derived that guarantee parameter convergence. Applications to parameter estimation in sigmoidal functions and identification of unknown frequencies of a sinusoidal signal are presented.

14. SUBJECT TERMS
Nonlinear parametrization, Triangular forms, Nonlinear systems, adaptive control, persistent excitation, parameter convergence, sigmoidal nonlinearities, frequency estimation

15. NUMBER OF PAGES

16. PRICE CODE

| 17. SECURITY CLASSIFICATION OR REPORT | 18. SECURITY CLASSIFICATION ON THIS PAGE | 19. SECURITY CLASSIFICATION OF ABSTRACT | 20. LIMITATION OF ABSTRACT |
|---|---|---|---|
| UNCLASSIFIED | UNCLASSIFIED | UNCLASSIFIED | UL |

NSN 7540-01-280-5500

Standard Form 298 (Rev.2-89)
Prescribed by ANSI Std. 239-18
298-102

Enclosure 1

20030605 067

# Adaptive Control of Nonlinearly Parametrized Systems
# DAAG55-98-1-0235

Anuradha M. Annaswamy
Final Report to the ARO



## Contents

# 1 Introduction

This project pertains to the development of a theory for adaptive control of nonlinear dynamic systems that are nonlinearly parametrized (NLP). Nonlinearly parameterized (NLP) dynamic systems are ubiquitous in practical applications. Problems related to friction compensation [1], reactors [2] are some of the examples. A large class of nonlinear models which are combinations of linear dynamics together with static nonlinearities such as Hammerstein-Uryson representations [3], as well as neural network models [4] are examples of dynamic systems with nonlinear parameterization.

One of the most common assumptions made in the context of adaptive control is that the unknown parameters occur linearly, and appear in linear [5] and nonlinear systems [6, 7, 8, 9]. The field of adaptive control has, by and large, treated the control problem in the presence of parametric uncertainties with the assumption that the unknown parameters occur linearly [5]. Whether in an adaptive observer or an adaptive controller, the assumption of linear parametrization has dictated the choice of the structure of the dynamic system. For instance, in adaptive observers, the focus has been on structures that allow measurable outputs to be expressed as linear, but unknown, combinations of accessible system variables. In direct adaptive control, a model-based controller is chosen so as to allow the desired closed-loop output to be linear in the control parameters. In indirect adaptive control, estimators and controllers are often chosen so as to retain the linearity in the parameters being estimated. The design of stable adaptive systems using the classical augmented approach as in [5] or using adaptive nonlinear designs as in [10] relies heavily on linear parametrization.

The problem is an important one both in linear and nonlinear dynamic systems, albeit for different reasons. In nonlinear dynamic systems, despite the fact that the majority of results have sought to extend the ideas of feedback linearization to systems with parametric uncertainties using the certainty equivalence principle, it is only within the context of linearly parametrizable nonlinear dynamics that global results have been available. Obviously, it is a nontrivial task to find transform methods for general nonlinear systems so as to convert them into systems with linear parametrizations. In linear dynamic systems, it is quite possible to transform the problem into one where unknown parameters occur linearly. However, such a transformation also can result in a large dimension of the space of linear parameters. This has a variety of consequences. The first is due to over parametrization which requires much larger amounts of persistent excitation or results in a lower degree of robustness. The second is that it can introduce undue restrictions in the allowable parametric uncertainty due to possible unstable pole-zero cancellations [11].

During the past three and half years, we have derived several new results related to NLP systems, and can be grouped under two categories: (i) Control of systems with a triangular structure[P1], and (ii) Parameter convergence in NLP systems[P2]-[P6]. These results are summarized below.

# 2 Triangular structures

The stage for controlling NLP systems has been set in [12, 13, 14, 15], where a new approach was developed to accommodate the parametric nonlinearity. In these investigations, it was assumed that the underlying system satisfies the matching conditions[16]. That is, the nonlinear system is of the form

$$\dot{X}_p = A_p X_p + b(f(\phi(t), \theta) + u) \tag{1}$$

where $f$ is a scalar nonlinearity in the unknown parameter $\theta$. This leads to the obvious question as to whther the approach in [12] can be extended to systems where these matching conditions are not satisfied. Investigations in this direction led to the following result:

The class of NLP-systems that we considered is of the form

$$\begin{aligned} \dot{z}_1 &= z_2 + f_1(z_1, \theta) \\ \dot{z}_2 &= z_3 \\ &\vdots \\ \dot{z}_n &= u \end{aligned} \tag{2}$$

where $\theta \in$ a compact set in $\mathbb{R}$. The goal is to find a $u$ that globally stabilizes the system when $\theta$ is unknown and $z_i$'s are accessible.

The following controller can be shown to lead to global stability. Define a tuning error $e_{n-1}$ using the following recursive relation:

$$\begin{aligned} e_0 &= z_1 \\ e_1 &= z_2 + g_1(z_1) \\ e_2 &= e_0 + z_3 + g_2(z_1, z_2) \\ &\vdots \\ e_i &= e_{i-2} + z_{i+1} + g_i(z_1, \ldots, z_i) \qquad i = 1, \ldots, n-1 \end{aligned} \tag{3}$$

where

$$\begin{aligned} g_i(z_1, \ldots, z_i) &= k_{i-1}(z_1, \ldots, z_{i-1}) + \overline{h}_{i-1}(z_1, \ldots, z_{i-1}) + \frac{\partial g_1}{\partial z_1} z_i, \qquad i = 2, \ldots, n-1 \\ k_0(z_1) = k_1(z_1) &= 0 \\ k_2(z_1, z_2) &= z_2 + \frac{\partial g_1^2}{\partial z_1^2} z_2 + \frac{\partial \overline{h}_1}{\partial z_1} z_2 \\ &\vdots \\ k_i(z_1, \ldots, z_i) &= k_{i-2} + z_i + \left( \frac{\partial g_1^2}{\partial z_1^2} z_i \right) z_2 + \frac{\partial g_1}{\partial z_1} z_{i-1} + \sum_{k=1}^{i-1} \left( \frac{\partial k_{i-1}}{\partial z_k} + \frac{\partial \overline{h}_{i-1}}{\partial z_k} \right) z_{k+1}, \\ &\quad i = 3, \ldots, n-1 \end{aligned}$$

$$\begin{aligned} h_0(z_1, \theta) = f_1(z_1, \theta); \qquad & h_1(z_1, \theta) = \frac{\partial g_1}{\partial z_1} f_1(z_1, \theta) \\ h_i(z_1, \ldots, z_i, \theta) &= h_{i-2} + \left( \frac{\partial g_1^2}{\partial z_1^2} z_i \right) z_i f_1(z_1, \theta) + \left( \frac{\partial k_{i-1}}{\partial z_1} + \frac{\partial \overline{h}_{i-1}}{\partial z_1} \right) f_1(z_1, \theta), \\ &\quad i = 2, \ldots, n-1 \end{aligned} \tag{4}$$

with $g_i(\cdot)$ and $\overline{h}_i(\cdot)$ such that

$$(f_1 - g_1)\sigma(e_0) \leq 0 \tag{5}$$
$$(h_i - \overline{h}_i)\sigma(e_i) \leq 0 \qquad i = 1, \ldots, n-1 \tag{6}$$

and the control input is given by

$$u = -\widehat{f}_n(z_1, z_2, \ldots, z_n) - \gamma_n e_{n-1} - e_{n-2} - a^* \mathrm{sat}(e_{n-1}) \tag{7}$$
$$\widehat{f}_n = k_{n-1}(z_1, \ldots, z_n) + \frac{\partial g_1}{\partial z_1} z_n + h_{n-1}(z_1, \ldots, z_n, \widehat{\theta}) \tag{8}$$
$$\dot{\widehat{\theta}} = -e_{n-1}\omega^* \qquad \gamma_n > 0 \tag{9}$$

$a^*$ and $\omega^*$ are solutions of the following min-max problem:

$$(a^*, \omega^*) = \min_{\omega} \max_{\theta} \left[ f_n - \widehat{f}_n + (\widehat{\theta} - \theta)\omega \right] \tag{10}$$

$g$ and $h$ can be chosen so as to satisfy eqs. (5) and (6) in the following manner: $g$ is a bounding function for $f$ chosen as

$$f(x_1) = \begin{cases} \max\limits_{\theta \in \Theta} [0, f(x_1, \theta)] + \epsilon_f & \forall\, x_1 > 0 \\ \min\limits_{\theta \in \Theta} [0, f(x_1, \theta)] - \epsilon_f & \forall\, x_1 < 0 \end{cases} \tag{11}$$

where $\epsilon_f$ is an arbitrary constant. Then, define a bounding function as:

$$G(x_1) = \begin{cases} F(x_1) & \forall\, |x_1| > \epsilon \\ S(x_1) & \forall\, |x_1| \leq \epsilon \end{cases} \tag{12}$$

where $\epsilon$ is also an arbitrary positive constant such that $\epsilon < \epsilon_f$, and $S(x_1)$ is an arbitrary smooth bounded function such that

$$\begin{aligned} S(x) &= F(x) \\ \frac{\partial S(x)}{\partial x} &= \frac{\partial F(x)}{\partial x} \end{aligned} \tag{13}$$

for $|x| = \epsilon$.

The proof of global stability of the proposed controller can be established as follows: From the definition of the errors $e_i$ in Eq. (3), it can be shown that

$$\begin{aligned} \dot{e}_0 &= e_1 + f_1 - g_1 \\ \dot{e}_1 &= z_3 + \frac{\partial g_1}{\partial z_1} z_2 + h_1 \end{aligned} \tag{14}$$

We define for $i = 1, \ldots, n$,

$$f_i(z_1, \ldots, z_i, \theta) = k_{i-1}(z_1, \ldots, z_{i-1}) + \frac{\partial g_1}{\partial z_1} z_i + h_{i-1}(z_1, \ldots, z_i, \theta)$$

and $z_{n+1} = u$. It can then be shown that

$$\dot{e}_1 = z_3 + f_2(z_1, z_2, \theta)$$

The method of induction can be used to show that

$$\dot{e}_i = z_{i+2} + f_{i+1}(z_1, \ldots, z_{i+1}, \theta) \tag{15}$$

This follows by using the relations

$$e_i = e_{i-2} + z_{i+1} + g_i$$

and

$$\dot{e}_{i-2} = z_i + f_{i-1}(z_1, z_2, \ldots, z_{i-1}, \theta)$$

and the definitions of $k_i$, $h_i$, and $f_i$. As a result equations (14), (15), $i = 1, \ldots, n-1$, and (7)-(9) define the closed-loop adaptive system.

We choose

$$V = \frac{1}{2}\left(\sum_{i=0}^{n-1} e_i^2 + \tilde{\theta}^2\right)$$

as a Lyapunov function candidate for the closed-loop adaptive system. This leads to the time-derivative

$$\begin{aligned} \dot{V} &= e_0\dot{e}_0 + \sum_{i=1}^{n-1} e_i(zi + 2 + f_{i+1}) - e_{n-1}\tilde{\theta}\omega^* \\ &= \sum_{i=0}^{n-2} e_i\left(f_{i+1} - g_{i+1}\right) + e_{n-1}\left(f_n - \widehat{f}_n - a^*\mathrm{sat}(e_{n-1}) + \tilde{\theta}\omega^*\right) \end{aligned} \tag{16}$$

Eqs. (5) and (6) imply that the first term is nonpositive. Eq. (10) implies that the second term is nonpositive. Together, they lead to the conclusion that $e_i$, $i = 0, \ldots, n-1$, and $\tilde{\theta}$ are bounded. From the definition of the $e_i$'s it can be shown that for $i = 1, \ldots, n$, $z_i$ and $u$ are bounded, and that $\lim_{t\to\infty} z_i(t) = 0$.

Extensions of the above result when (i) $z_1$ is required to track a desired trajectory $z_d$, (ii) when bounded disturbances $d_i$ are present, and (iii) for systems with triangular structures of the form

$$\begin{aligned} \dot{x}_1 &= \gamma_1(x_2) + f_1(z_1, \theta_1) \\ \dot{x}_2 &= \gamma_2(x)u + f_2(x_1, x_2, \theta_2) \end{aligned} \tag{17}$$

are currently being investigated. Preliminary investigations show that global boundedness can be established in each of these cases. A paper [P1] is currently under preparation summarizing all of the above results.

# 3 Parameter Convergence

A large class of problems in parameter estimation concerns nonlinearly parametrized systems (NLP). In the past few years, a stability framework for identification and control of such systems has been established. Over the past three years, we have addressed the issue of parameter convergence in such systems. Systems with both convex/concave and general parameterizations have been considered. The convergence results are stated separately below for discrete time systems and continuous time systems.

## 3.1 Convergence in discrete-time systems

In [15], we showed that for nonlinearly parameterized (NLP) discrete-time systems of the form of

$$y_t = f(\phi_{t-1}, \theta)$$

an adaptive estimator with a min-max algorithm leads to global stability. The question is whether the estimator can enable parameter convergence when the parameterization is concave or convex. We have derived sufficient conditions on the input $\phi$ and the nonlinearity $f$ under which parameter convergence results using the min-max algorithm. In section 3.1.2, a specific example of $f$ and $\phi$ that satisfy these conditions is presented.

### 3.1.1 Parameter convergence in the presence of concave/convex nonlinear parameterization

The underlying dynamic system is of the form:

$$\begin{aligned}
y_t &= f(\phi_{t-1}, \theta) \\
\widehat{y}_t &= f(\phi_{t-1}, \widehat{\theta}_{t-1}) \\
\widehat{\theta}_t &= \widehat{\theta}_{t-1} - \Gamma_\theta k_t \widetilde{y}_t \rho_t \omega_t \qquad \Gamma_\theta^T = \Gamma_\theta > 0 \\
k_t &= \frac{1}{\lambda + \omega_t^T \Gamma_\theta \omega_t} \qquad \lambda > 0 \\
\rho_t &= \max\{0, a_t\} \\
a_t &= 2 - \frac{2}{|\widetilde{y}_t|} J_0 \\
\omega_t &= arg \min_{\omega \in \mathbb{R}^n} \max_{\theta \in \Theta} \operatorname{sgn}(\widetilde{y}_t) J(\omega, \theta) \\
J(\omega, \theta) &= \widetilde{f}_{t-1} - \omega^T(\widehat{\theta}_{t-1} - \theta). \\
J_0 &= \min_{\omega \in \mathbb{R}^n} \max_{\theta \in \Theta} \operatorname{sgn}(\widetilde{y}_t) J(\omega, \theta)
\end{aligned} \tag{18}$$

where $\phi : N \to \mathbb{R}^n$. For any $\phi$ and all $\theta \in \Theta \subset \mathbb{R}^n$, where $\Theta$ is a compact set in $\mathbb{R}^n$, $f$ is assumed to be either concave or convex. The problem is to find conditions on $\phi_t$ under which $\widehat{\theta}_t$ converges to $\theta$ asymptotically.

It is assumed that the function $f$ at any time instant can be either concave or convex with respect to the parameter $\theta$. This property of $f$ shall be called as the curvature of $f$. It should be noted that the case when $f$ is linear in $\theta$ represents the transition between concavity and convexity or vice versa, and in such a case, the curvature can be labeled as either being convex or concave.

In LP systems, the term "persistently exciting" was used to characterize a signal which was rich enough in content to enable the convergence of parameter estimates to their true values by using the standard linear gradient-update algorithms (see [17, 5]). In order to distinguish it from its counterpart for LP systems, we will use the term "NLP persistent excitation" to specify a signal which allows convergence of parameter estimates to their true values in an NLP system, using the min-max algorithm. The required conditions for a signal to be NLP-persistently exciting are stated in Definition 1.

**Definition 1** *A function $\phi : N \to \mathbb{R}^n$ is NLP-persistently exciting with respect to $f(\phi, \theta)$, where $f : \mathbb{R}^n \times \Theta \to \mathbb{R}$, if at any time instant $t_a$, given any $\theta_1$, $\theta_2 \in \Theta$ and $\epsilon_\theta > 0$, such that $\|\theta_2 - \theta_1\| > \epsilon_\theta$, there exist positive constants $T$ and $\epsilon_f$ and a time instant $t_p \in [t_a + 1, t_a + T]$, such that*

*(NLP-I)* $\left|f\left(\phi_{t_p}, \theta_2\right) - f\left(\phi_{t_p}, \theta_1\right)\right| \geq \epsilon_f$;

**and** *at $t = t_p$,* **either**

*(NLP-IIa) sign* $\left(f(\phi_{t_p}, \theta_2) - f(\phi_{t_p}, \theta_1)\right) \neq$ *sign* $\left(f(\phi_{t_a}, \theta_2) - f(\phi_{t_a}, \theta_1)\right)$ *while the curvature of $f$ at $t_a$ and curvature of $f$ at $t_p$ are the same* **or**

*(NLP-IIb) sign* $\left(f(\phi_{t_p}, \theta_2) - f(\phi_{t_p}, \theta_1)\right) =$ *sign* $\left(f(\phi_{t_a}, \theta_2) - f(\phi_{t_a}, \theta_1)\right)$*, while the curvature of $f$ at $t_a$ and curvature of $f$ at $t_p$ are different.*

The requirements for NLP-persistent excitation consist of two components. The first component is condition *(NLP-I)* and, when $f$ is linear, it is equivalent to the LP persistent excitation, as shown below. Condition *(NLP-II)* is a second component of NLP-persistent excitation and is needed to overcome the presence of the dead-zone which in turn was required in the min-max algorithm to ensure stability. Condition *(NLP-II)* essentially states that, periodically, the probing input $\phi$ should be such that $f$ is appropriately dithered resulting in a change of either its curvature or its magnitude.

In order to establish parameter convergence in Eq. (18), we note first from the results of [15] that $\|\tilde{\theta}\|$ is non-increasing and that it is possible for adaptation to stop, which occurs whenever $\rho_t$ is small. To accommodate this behavior, the following notation is introduced. Let the set $\Omega_D$ denote the set of all time such that

$$\Omega_D = \{t \mid 0 \leq \rho_t < \epsilon_\rho\}, \qquad \text{where } \epsilon_\rho \text{ is a constant} \in (0, 2). \tag{19}$$

If $\epsilon_\rho$ is sufficiently close to zero, then $\Omega_D$ represents the time the system spends in the "dead-zone" where parameter adaptation is turned off. The complement of $\Omega_D$ is defined as $\Omega_D^C = \{t \mid \rho_t \geq \epsilon_\rho\}$. The question therefore is whether $\phi_t$ can be chosen so that the trajectories lie in $\Omega_D^C$ sufficiently often, which is answered in the affirmative below. As a first step, an important lemma which states a necessary condition for the system to be in the "dead-zone" is given. This is followed by Theorem 1 which presents the main result in parameter convergence.

**Lemma 1** *For the adaptive system given by Eq. (18), if $t \in \Omega_D$ then* **either**

*(D1)* $f_{t-1}$ *is concave in $\theta$,* and $\tilde{y}_t > 0$ **or**

*(D2)* $f_{t-1}$ *is convex in $\theta$,* and $\tilde{y}_t < 0$.

**Theorem 1** *For the system given by Eq. (18), if $\phi$ is NLP-persistently exciting and $\hat{\theta}_t \in \Theta \; \forall t$, then $\tilde{\theta}_t \to 0$ as $t \to \infty$.*

### 3.1.2 An example of NLP-persistent excitation

In this section, we provide one example of the function $f$ and a corresponding input $\phi$ that satisfy conditions *(NLP-I)* and *(NLP-II)* so as to enable parameter convergence. This example is given by:

$$f = e^{-\phi^T \theta} \tag{20}$$

where $\phi : N \to \mathbb{R}^n$, $\theta \in \Theta \subset \mathbb{R}^N$. The following definition states the desired property of the probing signal $\phi$.

**Definition 2** *Let $w \in \mathbb{R}^n$ be any unit vector. A bounded function $\phi : N \to \mathbb{R}^n$ is said to belong class $K^n$ if for any $t_a > t_0$, there exist positive constants $\epsilon_\phi$ and $T$, and a time instant $t_p \in [t_a + 1, t_a + T]$ such that*

$$\phi_{t_p}^T w \geq \epsilon_\phi$$

Definition 2 states that, periodically, the vector $\phi$ should have a positive component along every $w$ in $\mathbb{R}^n$. This is more restrictive than the linear case requirements [20], since the latter requires $\phi$ to merely have a nonzero component periodically along every vector in $\mathbb{R}^n$.

**Lemma 2** *For $f$ defined as in Eq. (20), $\phi \in K^n$ implies that $\phi$ is NLP-persistently exciting.*

An example of a function that belongs to class $K^2$ is

$$\phi = [\sin \nu t \, , \cos \nu t \,]^T \tag{21}$$

Since such a $\phi$ represents a rotating vector $\mathbb{R}^2$ with a constant angular velocity, it follows that it aligns itself with every $w$ in $\mathbb{R}^2$ periodically.

### 3.1.3 A special case

A specific example of an NLP system is one that contains sinusoidal signals with unknown frequencies and amplitudes, of the form

$$y(t) = \sum_{i=1}^{n} a_i \sin(\omega_i \phi(t))$$

and the task is to estimate $a_i$ and $\omega_i$ using all available data. Such problems arise in a number of applications related to vibration, noise suppression, and precise positioning [19]. Given that $\omega_i$ is a parameter that occurs nonlinearly, the question is whether direct estimation of $\omega_i$ and $a_i$ can be carried out using the approaches that were discussed above. Our preliminary investigations show that an estimation algorithm of the form

$$\begin{aligned} \widehat{y} &= \sum_{i=1} \widehat{a}_i \sin(\widehat{\omega}_i \phi) \\ \dot{\widehat{a}}_i &= -(\widehat{y} - y) \sin(\widehat{\omega}_i \phi) \\ \dot{\widehat{\omega}}_i &= -(\widehat{y} - y) g_i \end{aligned} \tag{22}$$

where $g_i$ correspond to the gradient of the nonlinearity evaluated at the current parameter estimate leads to global parameter convergence under certain conditions.

## 3.2 Parameter Convergence in Continuous-time Systems

In this section, results related to continuous-time systems are presented. First, we consider systems with convex/concave parameterization where sufficient conditions are derived under which parameter estimates converge to their true values using a min-max algorithm as in [12]. Next, we consider general parameterizatons where, to achieve parameter convergence, a hierarchical min-max algorithm is proposed where the lower-level consists of a min-max algorithm and the higher-level component updates the bounds on the parameter region within which the unknown parameter is known to lie. Using this hierarchical algorithm, a necessary and sufficient condition is established for global parameter convergence in systems with a general nonlinear parameterization. In both cases, the conditions needed are shown to be stronger than linear persistent excitation conditions that guarantee parameter convergence in linearly parametrized systems. Explanations and examples of these conditions and simulation results are included to illustrate the nature of these conditions are also included. A general definition of Nonlinear Persistent Excitation (NLPE) that leads to parameter convergence is proposed at the end of the paper.

Our objective is to identify unknown parameters in a class of nonlinear systems of the form

$$\dot{y} = -\alpha(u)y + f(\theta_0, u) \qquad 0 < \alpha_{min} \leq \alpha(u) \leq \alpha_{max}. \tag{23}$$

where $\theta_0 \in \Omega^0 \subset \mathbb{R}^n$ are bounded unknown parameters and $u \in \mathbb{R}^m$ are inputs. The function $f$ is a scalar valued function given by $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$.

We make the following assumptions regarding $u$ and $f$.

*Assumption 1* The input $u$ is Lipschitz in $t$ so that

$$\|u(t_1) - u(t_2)\| \leq U_b\|t_1 - t_2\|, \qquad \forall t_1, t_2 \in \mathbb{R}^+.$$

*Assumption 2* $f$ is Lipschitz with respect to its arguments, i.e.

$$|f(\theta + \Delta\theta, u + \Delta u) - f(\theta, u)| \leq B_\theta\|(\Delta u, \Delta\theta)\| \leq B_\theta(\|\Delta u\| + \|\Delta\theta\|).$$

Let a set $U_I$ be defined as follows:

$$U_I \;\; = \;\; \{u_i,\ i = 1, ..., I,\ u_i \neq u_j,\ i \neq j,\ \ u_i \in \mathbb{R}^m\}. \tag{24}$$

We introduce the definition of an identifiable function which is necessary for parameter convergence.

**Definition 3** *A function $f(\theta, u), \theta \in \Omega \subset \mathbb{R}^n$ is identifiable over parameter region $\Omega$ with respect to $U_I$ if there does not exist $\theta_1, \theta_2 \in \Omega$ and $\theta_1 \neq \theta_2$ such that*

$$\lim_{\theta \to \theta_1} f(\theta, u_i) = \lim_{\theta \to \theta_2} f(\theta, u_i) \qquad \forall u_i \in U_I, i = 1, ..., I.$$

Definition 3 implies that identifiability follows if the system of equations

$$f(\widehat{\theta}, u_i) - f(\theta_0, u_i) \;\; = \;\; 0 \qquad \forall u_i \in U_I \tag{25}$$

has a *unique* solution $\widehat{\theta} = \theta_0$ for any $\theta_0 \in \Omega$. Equation (25) provides a procedure for constructing $U_I$ such that for a given $\Omega$, $f$ can become identifiable over $\Omega$. That is, the number $I$ and the value $u_i$, for $i = 1, ..., I$ must be chosen such that Eq. (25) has a unique solution.

We also note that for a given $\Omega$, identifiability of $f$ is dependent on the choice of $U_I$. For example, if $f$ is linear, then $f$ is identifiable over any $\theta \in \mathbb{R}^n$ if elements of $U_I$ span the entire space of $\mathbb{R}^m$; for a nonlinear $f$, identifiability may be possible even if these elements span only a subspace. We notice that if $f$ is not identifiable with respect to $U_I$, it implies that we have no way of identifying $\theta_0$ using any input $u_i$ in $U_I$.

The dynamics of parameter estimation algorithm that we propose is the same as the min-max algorithm in [12] and is as follows: Suppose $\Omega^0$ is the unknown parameter region,

$$\begin{aligned} \dot{\tilde{y}} &= -\alpha(u)\tilde{y}_\epsilon + f(\hat{\theta}, u) - f(\theta_0, u) - a^* sat(\tfrac{\tilde{y}}{\epsilon}) \\ \dot{\tilde{\theta}} &= -\tilde{y}_\epsilon \phi^*. \end{aligned} \tag{26}$$

where

$$\tilde{y} = \hat{y} - y, \qquad \tilde{y}_\epsilon = \tilde{y} - \epsilon sat(\frac{\tilde{y}}{\epsilon}), \qquad \tilde{\theta} = \hat{\theta} - \theta_0, \tag{27}$$

and $a^*$ and $\phi^*$ come from the solution of an optimization problem

$$\begin{aligned} a^* &= \min_{\phi \in \mathbb{R}^m} \max_{\theta \in \Omega^k} g(\theta, u, \phi) \\ \phi^* &= arg \min_{\phi \in \mathbb{R}^m} \max_{\theta \in \Omega^k} g(\theta, u, \phi) \\ g(\theta, u, \phi) &= sat\left(\frac{\tilde{y}}{\epsilon}\right)\left(f(\hat{\theta}, u) - f(\theta, u) - \phi^T(\widehat{\theta} - \theta)\right). \end{aligned} \tag{28}$$

Our objective is to find conditions on $u$ and $f$ so that the system in (26) is uniformly asymptotically stable in the large.

### 3.2.1 Parameter Convergence in Systems with Convex/Concave Parameterization

The main stability result is stated in in this section in Theorem 2. As mentioned earlier, the region of attraction of the trajectories of (26) is shown to be a neighborhood $D_\epsilon$ of the origin $x = 0$.

**Theorem 2** *If (i) $f(\theta, u(t))$ is convex (or concave) on $\Omega$ for any $u(t) \in \mathbb{R}^m$, and (ii) for every $t_1 > t_0$, there exist positive constants $T_0$, $\epsilon_u$ and a time instant $t_2 \in [t_1,\ t_1 + T_0]$ such that for any $\theta$*

$$\beta(u(t_2))\left(f(\theta, u(t_2)) - f(\theta_0, u(t_2))\right) \geq \epsilon_u \|\theta - \theta_0\|, \tag{29}$$

*where $\beta(u_2) = 1$ if $f$ is convex and $-1$ is $f$ is concave, then all trajectories of (26) will converge uniformly to*

$$D_\epsilon = \{x \mid V(x) \leq \gamma_1\}, \tag{30}$$

*where*

$$\gamma_1 = \frac{2\epsilon}{\epsilon_u^2}\left(16 B_\theta U_b + 8 B_\theta B_\phi + 4 B_\phi^2\right), \tag{31}$$

*$\epsilon$ is defined as in (26), $\epsilon_u$ is given by (29), $U_b$ and $B_\theta$ are defined as in Assumptions 1 and 2, and $B_\phi$ is the bound on $\phi^*$ in (28) so that*

$$\|\phi^*(t)\| \leq B_\phi, \ \forall\, t \geq t_0, \qquad \forall\, t \geq t_0.$$

The proof of Theorem 2 follows by showing that if $u$ and $f$ are such that condition (29) is satisfied, then $\tilde{y}_\epsilon(t)$ becomes large at some time $t$ over the interval $[t_1, t_1 + T_0]$. Once $\tilde{y}_\epsilon(t)$ becomes large, it follows that $V(t)$ decreases over the interval $[t_1, t_1 + T_0]$ by a finite amount.

**Remark 1:** If $f$ is concave (or convex) for all $\theta \in \Omega$ and if $f$ satisifes the inequality in Eq. (29), we shall define that $f$ satisfies the Convex Persistent Excitation (CPE) condition with respect to $u$. Theorem 2 implies that if $f$ satisfies the CPE condition with respect to $u$, then parameter convergence to a desired precision $\epsilon$ follows.

**Remark 2:** From the definition of $D_\epsilon$, it automatically follows that as $\epsilon \to 0$, all trajectories converge to the region $x = 0$ and hence u.a.s.l. follows.

### 3.2.2 Sufficient Condition for Parameter Convergence

The CPE condition specifies certain requirements on $f$ in order to achieve parameter convergence. For a given $f$, theorem 2 does not state how $u$ should behave over time in order to satisfy (29). In this section, we state some observations and examples of $u$ that satisfies (29) for a general $f$.

Equation (29) consists of two separate requirements. Denoting $\tilde{f} = f(\theta, u) - f(\theta_0, u)$, the first requirement is that the magnitude of $\tilde{f}$ must be large. The second requirement is that $\tilde{f}$ must have the same sign as $\beta$. The first component states that for a large parameter error, there must be a large error in $\tilde{f}$. It is straightforward to demonstrate that this condition is equivalent to linear persistent excitation condition in [20], and is shown in section 3.2.3. The second requirement states what the sign of $\tilde{f}$ should be in relation to the convexity/concavity of $f$. If $f$ is convex, $\tilde{f}$ should be positive, and conversely, if $f$ is concave, $\tilde{f}$ should be negative.

The coupling of convexity/concavity and the sign of the integral of $\tilde{f}$ has the following practical implications. Suppose that $u$ is such that $f$ is always identifiable. To ensure parameter convergence, $u$ must be such that one of the following occurs: At least at one instant $t_2 \in [t_1,\ t_1 + T]$,

(a) For the given $\tilde{\theta}$, $u$ must change in such a way that the sign of $\tilde{f}$ is reversed, while keeping the convexity/concavity of $f$ the same, *or*

(b) For the given $\tilde{\theta}$, $u$ must reverse the convexity/concavity of $f$, while preserving the sign of $\tilde{f}$

### 3.2.3 Examples

We illustrate the above comments using specific examples of $f$. Suppose

$$f = e^{-u^T\theta} \tag{32}$$

where $u : \mathbb{R} \to \mathbb{R}^n$, $\theta \in \Omega \subset \mathbb{R}^n$. It can be checked that $f$ given in (32) is always convex with respect to $\theta$ for all $u$. Therefore, option *(b)* is not possible. Hence, $u$ must be such that $\tilde{f}$ can switch sign for any $\tilde{\theta}$ as required by option *(a)*. One example of such a $u$ is if for any $t_1$, there exists $t_2 \in [t_1,\ t_1 + T]$ such that

$$u^T(t_2)\, w \ \geq \epsilon_u \tag{33}$$

where $\omega$ is any unit vector in $\mathrm{IR}^n$. Another example which satisfies condition (29) is given by

$$f = \theta^u, \qquad u \in \mathrm{IR}.$$

It is easy to show that for such an $f$, condition (b) is satisfied if $u$ switches between $u_1$ and $u_2$ where $u_1 < 1$ and $u_2 > 1$.

The above examples show that the condition on $u$ that satisfies Eq. (29) varies with $f$.

**Relation to Conditions of Linear Persistent Excitation** The relation between CPE and LPE is worth exploring. For this purpose, we consider a linearly parameterized system, which is given by Eq. (23) with

$$f(\theta_0, u) = \theta_0^T \phi(u)$$

where $\phi(u) \in \mathrm{IR}^m$. In this case, it is well known that the corresponding estimator is given by equation (26) with $a^* = 0$ and $\phi^* = \phi$ [5]. The resulting error equations are summarized by

$$\begin{aligned} \dot{\tilde{y}} &= -\alpha(u)\tilde{y}_\epsilon + \tilde{\theta}^T \phi(u) \\ \dot{\tilde{\theta}} &= -\tilde{y}_\epsilon \phi(u). \end{aligned} \tag{34}$$

In [20], it is shown that u.a.s.l. of (34) follows under an LPE condition. For the sake of completeness, we state this condition below.

**Definition (LPE):** $\phi$ is said to be linearly persistent exciting (l.p.e.) if for every $t_1 > t_0$, there exists positive constants $T_0$, $\delta_0$, $\epsilon_0$ and a subinterval $[t_2,\ t_2 + \delta_0] \in [t_1,\ t_1 + T_0]$ such that

$$\left| \int_{t_2}^{t_2+\delta_0} \theta^T \phi(u(\tau)) d\tau \right| \geq \epsilon_0 \|\theta\|. \tag{35}$$

We now show the relation between the LPE condition and the CPE condition in (29). When $f(\theta, u) = \theta^T \phi(u)$, if Assumption 1 holds, it can be shown that the LPE condition is equivalent to the following inequality: For every $t_1 > t_0$ and $\theta$, there exists positive constants $T_0$, $\epsilon_0$ and a time instant $t_2 \in [t_1,\ t_1 + T_0]$ such that

$$|f(\theta, u(t_2)) - f(\theta_0, u(t_2))| \geq \epsilon_u \|\hat{\theta} - \theta_0\|. \tag{36}$$

Since a linear function can be considered to be either convex or concave, the inequality in (36) is equivalent to the CPE condition in (29). This equivalence is summarized in the following lemma:

**Lemma 3** *When $f(\theta, u) = \theta^T \phi(u)$, if Assumption 1 holds, the CPE condition in Eq. (29) is equivalent to the LPE condition in Eq. (35).*

It should be noted that for a general nonlinear $f$, the CPE condition becomes more restrictive than the LPE condition. For example, for $f$ as in (32), the CPE condition implies that $u$ must satisfy (33). On the other hand, if $f = u^T\theta$, even if $u$ is such that $|u^T(t_2)w|$ is periodically large, the LPE condition is satisfied.

**A Counter-example** For a general function $f$, it may not be possible to find a $u$ that satisfies either condition (a) or (b) mentioned above. A simple example is

$$f = cos(\theta u)$$

where $|u| \leq u_{max}$ and $\theta \in [0,\ \pi/(2u_{max})]$. We note that $f$ is concave and monotonically decreasing for any $u$ with $|u| \leq u_{max}$. Hence neither (a) nor (b) is satisfied. That is, it is possible for the min-max algorithm to result in the parameter estimate $\hat{\theta}$ to get "stalled" in a region in$\Omega^0$. This motivates the need for an improved min-max algorithm, and is outlined in section 3.2.4.

### 3.2.4 Parameter Convergence in Systems with a General Parameterization

In the previous section, we showed that if a function $f$ is convex (or concave), and if $f$ and $u$ satisfy the CPE condition, then parameter convergence follows. However, as we saw in section 3.2.3, not all convex/concave functions can satisfy the CPE condition. In this section, we present a new algorithm which not only allows the persistent excitation condition to be relaxed but also enables parameter convergence for non-convex and non-concave functions.

The algorithm we present in this section is hierarchical in nature, and consists of a lower-level and a higher-level. In the lower-level, for a given unknown parameter region $\Omega^0$, the parameter estimate $\hat{\theta}$ is updated using the min-max algorithm as in (26). In the higher-level, using information regarding the parameter estimate $\hat{\theta}$ obtained from the lower-level, the unknown parameter region is updated as $\Omega^1$. Iterating between the lower and higher levels, the overall hierarchical algorithm guarantees a sequence of parameter region $\Omega^k$. The properties of these two levels are discussed below.

**Lower-level Algorithm** The lower-level algorithm consists of the min-max parameter estimation as in (26) with the unknown parameter $\theta_0 \in \Omega^k$. We show in this section that when this algorithm is used, the asymptotic convergence of $\tilde{y}_\epsilon(t)$ to zero occurs in a finite time. Once $\tilde{y}_\epsilon$ becomes small, we estimate the region that the unknown parameter can lie in using the corresponding parameter estimate $\hat{\theta}^k$. This region is used in the higher-level part of hierarchical algorithm to update the unknown parameter region from $\Omega^k$ to $\Omega^{k+1}$. The convergence of $\tilde{y}_\epsilon$ is stated in Lemma 4, and the characterization of the unknown parameter is stated in Lemma 5.

**Lemma 4** *For the system in (23) and the estimator in (26), given any positive $T$ and $\delta$, there exists a finite time $t_1$ such that*

$$|\tilde{y}_\epsilon(t)| \leq \delta \qquad \text{for } t_1 \leq t \leq t_1 + T. \tag{37}$$

We note that for every specific input $u$, a time $t_1$ that satisfies (37) exists. However the value of $t_1$ will depend on the choice odf $u$. Since our goal is parameter convergence, we require $u$ to assume distinct values, i.e. persistently span a set of interest. This is stated in the definition below.

Let $U_I$ be defined as in Eq. (24).

**Definition 4** *$u$ is said to persistently span $U_I$ if for any $u_i \in U_I$ and any $t_1$, there exist a finite $T_i$ and $\tau_i$ such that*

$$u(\tau_i) = u_i \qquad \tau_i \in [t_1,\ t_1 + T_i] \qquad i = 1, ..., I. \tag{38}$$

Definition 4 implies that $u$ periodically visits all points in $U_I$.

Let

$$B_t = 2B_\theta(\delta B_\phi + 2U_b) + \delta B_\phi^2, \tag{39}$$

if we choose $T$ as

$$T = \max_{1 \leq i \leq I} T_i + \frac{2\sqrt{B_t(\delta + \epsilon)}}{B_t}$$

where $T_i$ is given by (38), then Lemma 4 implies that there exists a finite time $t_1$ such that

$$|\tilde{y}_\epsilon(t)| \leq \delta \qquad t_1 \leq t \leq t_1 + T. \tag{40}$$

When $\tilde{y}_\epsilon$ satisfies (40), we refer to it as lower-level convergence. If $u$ persistently spans $U_I$, then Definition 1 and the choice of $\tau_i$ implies that at $\tau_i \in [t_1,\ t_1 + T]$, $u(\tau_i) = u_i$, $i = 1, ..., I$. The parameter estimate $\hat{\theta}(\tau_i)$ at time instances are defined as

$$\hat{\theta}_i^c = \hat{\theta}(\tau_i) \qquad i = 1, ..., I,$$

and are denoted as low-level convergent estimates. We characterize the region where the unknown parameters lie in lemma 5 using these lower-level convergent estimates.

**Lemma 5** *For the system in (23) and estimator in (26), let $\Omega$ be the unknown parameter region and $\hat{\theta}_i^c, i = 1, ..., I$, be the lower-level convergent estimates. If the input $u$ persistently spans $U_I$, then*

$$\theta_0 \in \bigcap_{i=1}^{I} \Phi_\epsilon(\Omega, u_i, \epsilon, \delta, \hat{\theta}_i^c).$$

*where*

$$\begin{aligned}
\Phi_\epsilon(\Omega, u_i, \epsilon, \delta, \hat{\theta}_i^c) &= \left\{\theta \in \Omega | \underline{f}_i \leq f(\theta, u_i) \leq \overline{f}_i\right\} \\
\underline{f}_i &= f(\hat{\theta}_i^c, u_i) - a_+^*(\hat{\theta}_i^c, u) - \alpha_{max}\delta - 2\sqrt{B_t(\delta + \epsilon)} \\
\overline{f}_i &= f(\hat{\theta}_i^c, u_i) + a_-^*(\hat{\theta}_i^c, u) + \alpha_{max}\delta + 2\sqrt{B_t(\delta + \epsilon)}\}
\end{aligned}$$

*and $B_t$ as in (39).*

**Higher-Level Algorithm** We now present the higher-level component of the hierarchical algorithm. Here, our goal is to start from a known parameter region $\Omega^k$ that the unknown parameter $\theta_0$ lies in, and update it as $\Omega^{k+1}$ using all available information from the lower-level component. In particular, we use $\Phi_\epsilon$ to update $\Omega^k$. In order to reduce the parameter uncertainty, different $\Phi_\epsilon'$s are computed by varying $u_i$, $i = 1, ..., I$. The resulting $\Omega^{k+1}$ is chosen as

$$\Omega^{k+1} = \bigcap_{i=1}^{I} \Phi_\epsilon(\Omega^k, u_i, \epsilon, \delta, \hat{\theta}_i^c).$$

We state the complete hierarchical algorithm below:

Step H1: Set $k = 0$ and $\Omega^k = \Omega^0$, and $T = \max_i T_i + 2\sqrt{B_t(\delta + \epsilon)}/B_t$.

Step H2: Run the estimator in (26). Wait until time $t_k^*$ where

$$|\tilde{y}_\epsilon(t)| \leq \delta \qquad \text{for } t \in [t_k^*,\ t_k^* + T],$$

and record the low level convergent estimate $\hat{\theta}_i^c$ as

$$\hat{\theta}_i^c = \hat{\theta}(\tau_i)$$

where

$$u(\tau_i) = u_i, \qquad \forall\ \tau_i \in [t_k^*,\ t_k^* + T].$$

Step H3: Calculate $\Omega^{k+1}$ from $\Omega^k$ and $\hat{\theta}_i^c, i = 1, ...I$, as follows:

$$\Omega^{k+1} = \bigcap_{i=1}^{I} \Phi_\epsilon(\Omega^k, u_i, \epsilon, \delta, \hat{\theta}_i^c).$$

Step H4: If $\Omega^{k+1} = \Omega^k$, stop. Otherwise, set $k = k + 1$ and return to step 2.

The question that remains to be answered is whether $\Omega^{k+1}$ is a strict subset of $\Omega^k$ so that parameter convergence of $\hat{\theta}$ to $\theta_0$ can be ensured.

**Parameter Convergence with the Hierarchical Algorithm** We now address the question of parameter convergence of the hierarchical algorithm. For ease of exposition, we set $\epsilon = \delta = 0$. The effects of nonzero $\epsilon$ and $\delta$, as mentioned above, simply affect the accuracy of the parameter error. For $\epsilon = \delta = 0$, the hierarchical algorithm can be summarized as follows:

Step 1: Set $k = 0$ and $\Omega^k = \Omega^0$, and $\delta \to 0$.

Step 2: Run the estimator in (26). Wait until time $t_k^*$ where

$$|\tilde{y}_\epsilon(t)| \leq \delta \qquad \text{for } t \in [t_k^*,\ t_k^* + T],$$

and record the low level convergent estimate $\hat{\theta}_k^c = \hat{\theta}(\tau) \;\; \tau \in [t_k^*,\ t_k^* + T]$

Step 3: Calculate $\Omega^{k+1}$ from $\Omega^k$ and $\hat{\theta}_k^c$ as follows:

$$\Omega^{k+1} = \{\theta \in \Omega^k \mid \underline{f}_i^k \leq f(\theta_0, u_i) \leq \overline{f}_i^k, i = 1, ..I.\} \tag{41}$$

where

$$\begin{aligned} \underline{f}_i^k &= f(\hat{\theta}_k^c, u_i) - a_+^*(\hat{\theta}_k^c, u_i), \qquad & i = 1, ...I \\ \overline{f}_i^k &= f(\hat{\theta}_k^c, u_i) + a_-^*(\hat{\theta}_k^c, u_i), \qquad & i = 1, ...I \end{aligned}$$

Step 4: If $\Omega^{k+1}$ reduces to the desired precision, we are done. Otherwise, set $k = k + 1$ and return to step 2.

We first define a measure $\widehat{V}(\Omega^k)$ as

$$\widehat{V}(\Omega^k) = \sum_{i=1}^{m+1} (\overline{f}_i^k - \underline{f}_i^k)$$

where $\overline{f}_i^k$ and $\underline{f}_i^k$ are defined as in (41). Next, we introduce a definition for a "stalled" parameter region $\Delta_i$:

*Definition 3:* For any $\Omega \subseteq \Omega_0$, define $\underline{f}_i^*$ and $\overline{f}_i^*$ as

$$\underline{f}_i^* = \min_{\theta \in \Omega^*} f(\theta, u_i), \qquad \overline{f}_i^* = \max_{\theta \in \Omega^*} f(\theta, u_i). \tag{42}$$

Then we define $\Delta_i(\Omega)$ to be a "stalled" estimate-region of $\Omega$ as

$$\Delta_i(\Omega) = \{\theta \mid f(\theta, u_i) - a_+^*(\theta, u_i) \leq \underline{f}_i^* \text{ and } f(\theta, u_i) + a_-^*(\theta, u_i) \geq \overline{f}_i^*\}. \tag{43}$$

We now prove a property of $\Delta_i(\Omega)$ which explains why it corresponds to a "stalled" region in $\Omega$.

**Lemma 6** *If $\hat{\theta}^c \in \Delta_i(\Omega^k)$ for any $k$, then*

$$\Omega^j = \Omega^k, \qquad j = k+1, \ldots$$

Proof of Lemma 6 follows directly from the definition of $\Delta_i$ and the construction of $\Omega^{k+1}$ in (41).

In Theorem 3, we propose the conditions needed for the higher-level convergence.

**Theorem 3** *For the system in (23) and estimator in (26), let $f$ be identifiable over $\Omega$ w.r.t. $U_I$ and $u$ persistently span $U_I$. The hierarchical algorithm outlined in Steps H1-H4 guarantees that*

$$\lim_{t \to \infty, \delta \to 0, \epsilon \to 0} \hat{\theta}(t) = \theta_0 \tag{44}$$

*if and only if for any $\Omega \subseteq \Omega_0$ where $\theta_0 \in \Omega$,*

$$\bigcap_{i=1,\ldots,I} \Delta_i(\Omega) = \phi \text{ or } \bigcap_{i=1,\ldots,I} \Delta_i(\Omega) = \{\theta_0\}. \tag{45}$$

*where $\phi$ denotes the null set.*

**Remark 3:** If $f(\theta, u)$ is identifiable over $\Omega$ with respect to $U_I$, $u$ persistently spans $U_I$, and $f$ satisfies the inequality (45), we shall define that $f$ satisfies the Nonlinear Persistent Excitation (NLPE) condition with respect to $u$. Theorem 2 implies that then NLPE of $f$ with respect to $u$ is necessary and sufficient for parameter convergence to take place.

**Remark 4:** The requirement on $u$ for $f$ to satisfy the NLPE can sometimes be less stringent than that on $u$ for LPE. An example of this fact is for the parameter $\theta = [a, \omega]^T$, and the cases (i) $l(\theta) = \theta^T u$, and (ii) $f(\theta) = a u_1 \cos(\omega u_2)$ where $u_1$ and $u_2$ are the elements of $u$. Clearly, for a $u$ such that $u_1 = k u_2$, where $k$ is a constant, $u$ does not satisfy $LPE$, but $f$ does satsify NLPE with respect to $u$. As shown in Section 3.2.3, NLPE can impose more stringent conditions on $u$ as well.

**Remark 5:** It should be noted that the NLPE condition guarantees parameter convergence for any general nonlinear function $f$ that is identifiable. This implies that the min-max algorithm outlined in [21], which is applicable for even a non-convex (or a non-concave) function, can be used to establish parameter convergence. We include simulation results of such an example in Section 3.2.5.

**Remark 6:** It should be noted that a fairly extensive treatment of conditions of persistent excitation has been carried out in [22, 23] for a class of nonlinear systems. The systems under consideration in this paper do not belong to this class. The most distinct features of the system (23) is the presence of the quantity $a^*$ and the quantity $f(\widehat{\theta}, u) - f(\theta_0, u)$, where the former can introduce equilibrium points other than zero and the latter is not Lipschitz with respect to $\widehat{\theta} - \theta$. As a result, an entirely different set of conditions and properties have had to be derived to establish parameter convergence.

**Relation between NLPE and CPE** In what follows we compare the NLPE and the CPE conditions. In order to facilitate this comparison, we restate the CPE condition in a simpler form:

**Definition 5** *$f$ is said to satisfy the CPE′ condition with respect to $u$ if (i) $f(\theta, u(t))$ is convex (or concave) on $\Omega$ for any $u(t) \in \mathbb{R}^m$, and (ii) $u$ is persistently spanning with respect to $U_I$, and*

$$\text{(iii)} \qquad \beta(u_i)\left(f(\theta, u_i) - f(\theta_0, u_i)\right) \geq \epsilon_u \|\theta - \theta_0\|, \qquad \forall u_i \in U_I, \tag{46}$$

We note that the only distinction between the inequalities in (29) and (46) is in the value taken by $u(t_2)$ for some $t_2$ in the interval $[t, t+T]$. In (46) it implies that $u(t_2)$ assumes one of the finite values $u_i$ in $U_I$ while in (29), the corresponding $U_I$ can consist of infinite values. If $u$ is "ergodic" in nature so that it visits all typical values that it will assume for all $t$ over one interval, then it implies that the two conditions (29) and (46) are equivalent. We shall assume in the following that the input is "ergodic."

**Lemma 7** *Let $f(\theta, u_i)$ be convex (or concave) for all $\theta \in \Omega^0$ and $\Omega \subseteq \Omega^0$. Then the CPE′ condition implies the NLPE condition.*

**Remark 5:** Lemma 7 shows that the CPE′ condition is sufficient for the NLPE to hold if $f$ is convex (or concave). Clearly, the CPE′ condition is not necessary, as shown by the counterexample in Section 3.2.3. The NLPE condition therefore represents the most general definition of persistent excitation in nonlinearly parameterized systems.

### 3.2.5 Simulation Results

We consider the system in (23) and the estimator in (26) to evaluate the performance of the hierarchical algorithm. The system parameters are chosen as follows:

$$f = \left(\theta - \frac{u}{8}\right)^2 + 12exp\left\{-5\left(\theta - 2 + \frac{u}{4}\right)^2\right\}$$

where $\theta$ is an unknown parameter that belongs to a known interval $\Omega = [0, 5]$. The input $u$ is chosen as a sinusoidal function $u = 1.1\sin(2t)$. We note that the function $f$ is non-convex (and non-concave), whose values are shown in Figure 2 for $u = 1, -1, 0$. It can be shown that $f$ is identifable with respect to $\Omega$ and that $u$ is persistently spanning with respect to $U_I = \{1, -1, 0\}$.

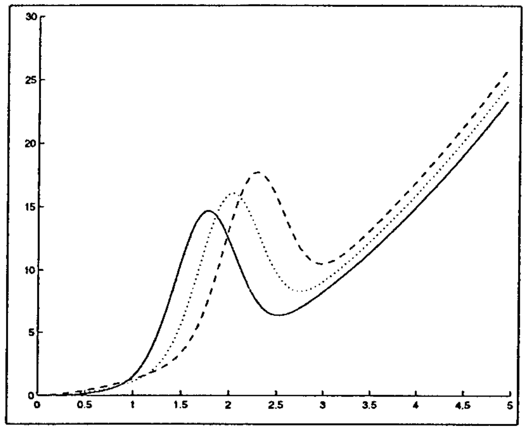


Figure 1. A non-concave (and non-convex) function $f(\theta, u)$ vs. $\theta$, for $u = 1, 0, -1$. $f(\theta, 1)$:—, $f(\theta, -1)$:- - -, $f(\theta, 0)$:.....

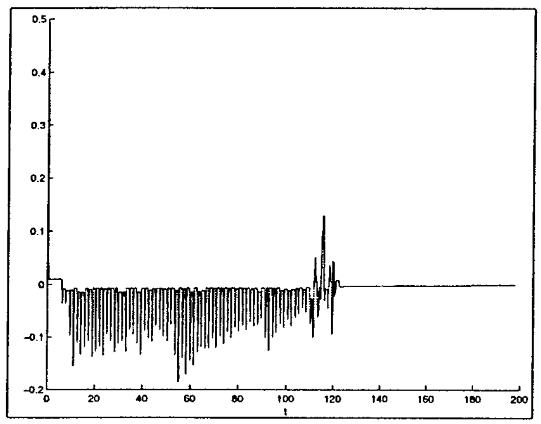


Figure 2. The output error $\tilde{y}_\epsilon(t)$ with $t$ using the hierarchical algorithm. $\epsilon = 0.001$ and $\delta = 0.02$.

The hierarchical algorithm in (26) together with steps H1 to H4 was implemented to estimate $\theta$. The parameters $\epsilon = 0.001$ and $\delta = 0.02$. Since $u$ is a sinusoid, the parameter $T$ was set to the corresponding period $\pi$. The resulting output error $\tilde{y}_\epsilon$, parameter estimate $\hat{\theta}$, and the update of the parameter region $\Omega$ are shown in Figures 3- 5, respectively. The evolutions of the lower and upper bounds $\underline{f}_i^k$ and $\overline{f}_i^k$, $i = 1, 2, 3$ with respect to $t$ are also shown in Figure 6. A similar convergence was observed to occur for any $\theta_0$ in $\Omega$. These figures show that the update of $\Omega^k$ is not necessarily periodic. Once $\tilde{y}_\epsilon$ becomes smaller than $\delta$ over an interval $T$, the corresponding parameter estimates and the upper and lower bounds on $f_i$ and therefore on $\Omega$ are computed. It was also observed that just the min-max algorithm without the higher level component did not result in parameter convergence.

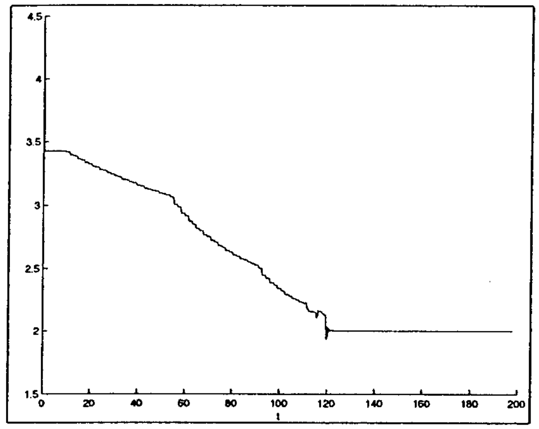

Figure 3. The parameter estimate $\widehat{\theta}(t)$ with $t$ using the hierarchical algorithm. True parameter value $\theta_0 = 2$.

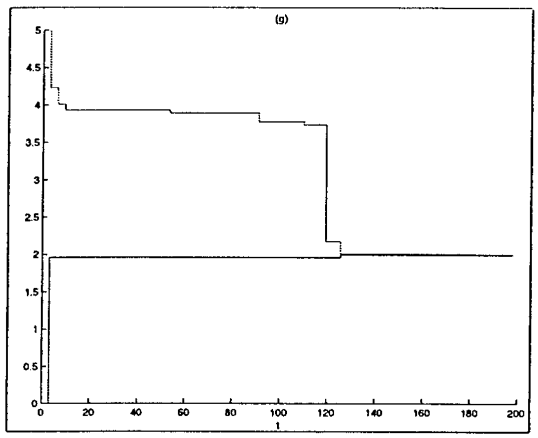

Figure 4. The evolution of the parameter region $\Omega^k$ with $t$, using the hierarchical algorithm. Note that $\Omega^k$ is updated at instants $t_k^*$ such that $|\tilde{y}_\epsilon(t)| \leq \delta$ for $t \in [t_k^* - T, t_k^*]$.

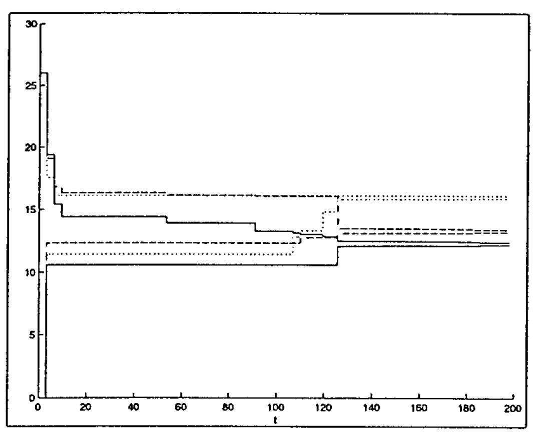

Figure 5. The upper-bounds $\overline{f}_i^k$ and lower-bounds $\underline{f}_i^k$ of $f(\theta, u_i)$ with $t$ using the hierarchical algorithm, for $u_i = 1, -1, 0$. $\overline{f}_1$,$\underline{f}_1$:—, $\overline{f}_2$,$\underline{f}_2$:- - -, $\overline{f}_2$,$\underline{f}_2$:....

## 4 Summary

Under this award, a systematic adaptive control theory for nonlinearly parameterized systems has been derived. Adaptive control results for systems with a triangular structure have been demonstrated, establishing global stability. Significant results for parameter convergence in such systems have been derived, establishing the concept of *Nonlinear Persistent Excitation* for systems with nonlinear parameterization and are illustrated using several examples. These results are expected to have a significant impact on modeling and control of complex physical systems.

## 5 Personnel and Inventions

The various personnel who have been supported through this award are listed below:

1. Dr. Anuradha Annaswamy
2. Dr. Aleksandar Kojic, received his PhD in September 2001.
3. Chengyu Cao, PhD Candidate

No inventions were developed through the support of this grant.

## 6 Publications

[P1 ] A. Kojic and A.M. Annaswamy, "Adaptive control of nonlinearly parameterized systems with a Triangular Structure," *Automatica* 38:115-123, 2002.

[P2 ] F.P.Skantze, A.Kojic, A.P.Loh and A.M.Annaswamy, "Adaptive estimation of discrete-time systems with nonlinear parameterization" *Automatica* 36(12):1879-1887, 2000.

[P3 ] A.Kojic, C.Cao and A.M.Annaswamy, "Parameter convergence in systems with convex/concave parameterization", *Proceedings of the American Control Conference*, Chicago, IL, pp.2240-2244, 2000.

[P4 ] A.M.Annaswamy, N.T.Ho, C.Cao, and A.Kojic, "A convergent Frequency estimator", *Proceedings of the American Control Conference*, Chicago, IL, pp.2235-2239, 2000.

[P5 ] C.Cao, and A.M.Annaswamy, "Parameter Convergence in systems with a general nonlinear parameterization using a hierarchical algorithm", *Proceedings of the American Control Conference* (to appear), 2002.

[P6 ] C.Cao, A.Kojic and A.M.Annaswamy, "Parameter Convergence in nonlinearly parameterized systems", submitted to *IEEE Transactions on Automatic Control*, December 2001.

# References


[1] B. Armstrong-Hélouvry. *Control of Machines with Friction*. Kluwer Academic Publishers, Norwell, MA, 1991.

[2] Jovan D. Bosković. Stable adaptive control of a class of fed-batch fermentation processes. In *Proceedings of the Eighth Yale Workshop on Adaptive and Learning Systems*, pages 263–268, Center for Systems Science, Dept. of Electrical Engineering, Yale University, New Haven, CT 06520-1968, June 1994.

[3] Esref Eskinat, Stanley H. Johnson, and William L Luyben. Use of hammerstein models in identification of nonlinear systems. *American Institute of Chemical Engineers journal*, 37(2):255–268, February 1991.

[4] Kumpati S. Narendra and Kannan Parthasarathy. Identification and control of dynamical systems using neural networks. *IEEE Transactions on Neural Networks*, 1(1):4–26, March 1990.

[5] K. S. Narendra and A. M. Annaswamy. *Stable Adaptive Systems*. Prentice-Hall, Inc., Englewood Cliffs, N.J., 1989.

[6] S. S. Sastry and M. Bodson. *Adaptive Control: Stability, Convergence, and Robustness*. Prentice-Hall, Inc., Englewood Cliffs, N.J., 1989.

[7] I. Kanellakopoulos, P.V. Kokotovic, and A.S. Morse. Systematic design of adaptive controllers for feedback linearizable systems. *IEEE Transactions on Automatic Control*, 34:1241–1253, November 1991.

[8] M. Krstić, I. Kanellakopoulos, and P. V. Kokotović. Adaptive nonlinear control without overparameterization. *Systems & Control Letters*, 19:177–186, March 1992.

[9] Danbing Seto, Anuradha M. Annaswamy, and J. Baillieul. Adaptive control of nonlinear systems with a triangular structure. *IEEE Transactions on Automatic Control*, 39(7):1411–1428, July 1994.

[10] M. Krstić, I. Kanellakopoulos, and P. V. Kokotović. Nonlinear design of adaptive controllers for linear systems. *IEEE Transactions on Automatic Control*, 39(4):738–752, April 1994.

[11] A. S. Morse. Supervisory control of families of linear set-point controllers - Part I: Exact matching". Technical Report 9301, Yale University, March 1993.

[12] A.M. Annaswamy, A.P. Loh, and F.P. Skantze. Adaptive control of continuous time systems with convex/concave parametrization. *Automatica*, 34:33–49, January 1998.

[13] A.M. Annaswamy, N. Mehta, C. Thanomsat, and A.P. Loh. Application of adaptive controllers based on nonlinear parametrization to chemical reactors and magnetic bearings". In *Proceedings of the ASME Dynamic Systems and Control Division*, Dallas, TX, November 1997.

[14] A. P. Loh, A.M. Annaswamy, and F.P. Skantze. Adaptation in the presence of a general nonlinear parametrization: An error model approach. *IEEE Transactions on Automatic Control*, (accepted for publication) 1998.

[15] F.P. Skantze, A.P. Loh, and A.M. Annaswamy. Adaptive estimation of discrete-time systems with nonlinear parametrization. *Automatica*, (submitted) 1997.

[16] D.G. Taylor, P.V. Kokotovic, R. Marino, and I. Kanellakopoulos. "Adaptive regulation of nonlinear systems with unmodeled dynamics". *IEEE Transactions on Automatic Control*, 34:405–412, 1989.

[17] G. C. Goodwin and K. S. Sin. *Adaptive Filtering, Prediction, and Control*. Prentice-Hall, 1984.

[18] B.D.O. Anderson and Jr. C.R. Johnson. "Exponential convergence of adaptive identification and control". *Automatica*, vol. 18:1–13, 1982.

[19] C. Canudas de Wit and L. Praly. Adaptive eccentricity compensation. In *Proceedings of the Conference on Decision and Control*, Tampa, FL, 1998.

[20] A.P. Morgan and K.S. Narendra. "On the stability of nonautonomous differential equations $\dot{x} = [A + B(t)]x$ with skew-symmetric matrix $B(t)$". *SIAM Journal of Control and Optimization*, vol. 15:163–176, January 1977.

[21] A. P. Loh, A.M. Annaswamy, and F.P. Skantze. Adaptation in the presence of a general nonlinear parametrization: An error model approach. *IEEE Transactions on Automatic Control*, AC-44:1634–1652, September 1999.

[22] E. Panteley, A. Loria, and A. Teel. Relaxed persistency of excitation for uniform asymptotic stability. *IEEE Transactions on Automatic Control*, (to appear), 2001.

[23] Y. Zhang, P. Ioannou, and C. Chien. Parameter convergence of a new class of adaptive controllers. *IEEE Transactions on Automatic Control*, 41:1489–1493, 1996.